



Summary

Main Features

Census Working Paper 94/2

1991 CENSUS DATA QUALITY: EDUCATION

Julie Evans
Darren McDonald

Population Census Evaluation
October 1994

CONTENTS

Introduction

Attendance at an Educational Institution

- Background
- Quality of Responses
- Comparison with non-census data
- Summary

Age Left School

- Background
- Quality of responses
- Comparison with NSSC data for level of Schooling
- Cohort analysis, 1986 and 1991 Censuses
- Summary

Educational Qualifications

- Background
- Analysis of sequencing question
- Analysis of Field of Study and Level of Attainment
- Analysis of year of Qualification
- Effectiveness of the 'Still studying for first qualification' response category

INTRODUCTION

The demand for data on education is derived in part from a wide range of socio-economic research activities that are conducted by both government and non-government organisations.

The governments of all Australian States and Territories and the Commonwealth have a major responsibility for funding and administering education at primary, secondary and tertiary levels. Conversely, funding education is a major component of public outlays. Data is needed by governments as part of the process of managing and evaluating their education resources and programs.

Education data can be used to study educationally disadvantaged groups, relate occupation, industry and income data, evaluate the levels of skills in the labour force and develop labour force training programs. In the 1991 Census, there were eight questions covering three topics about education: Attendance at an Educational Institution, Age Left School and Educational Qualifications.

The Census questions on Attendance at an Educational Institution are designed to measure current levels of participation in education and identify which type of educational institutions are being attended. This information can be used for education policy making decisions such as planning facilities and allocating resources for education programs.

In recent years, governments have been addressing the demand for more skilled labour in the high technology industries by encouraging higher retention rates in Years 11 and 12 and greater participation levels in mathematics and science at tertiary level. Census information can be used to measure the impact of these policies on enrolment patterns as well as the overall demand for education services.

Education data can be used to augment the analysis of other demographic statistics such as health, employment and economic data. Age Left School, for example, is designed to identify the highest level of primary or secondary school completed and is used as an indicator of socio-economic disadvantage for a range of government funding programs.

Data on education qualifications are important for the assessment of the skill level of the labour force and are therefore valuable for the planning and implementation of labour force training programs. The questions on Educational Qualifications are designed to identify details of the highest qualification obtained by a person. The Department of Employment, Education and Training uses the data as a basis for the preparation of the 'Occupational Demand Schedule' for use by the Department of Immigration, Local Government and Ethnic Affairs in recruiting skilled migrants.

Developments in the wider education structure and the classification of education data had significant implications for the 1991 Census.

The restructuring within the tertiary education system that occurred in the late 80's resulted in a larger diversity of courses offered by institutions, and made it increasingly difficult to label institution type. A number of variations to the form design were made to the educational attendance question for the 1991 Census in response to the changing circumstances.

In 1991 a new classification for education qualifications, the ABS Classification of Qualifications (ABSCQ), was introduced. ABSCQ was used to code Census qualifications

data and help standardise the different qualifications recorded.

These and other developments are examined more closely to determine what effects they have had on Census data. The aim of this paper is to provide users with a detailed assessment of the benefits and limitations of Census data for each of the three education topics contained in the 1991 Census. Each topic is analysed separately looking, in particular, at factors such as variation in counts between censuses, changes in form design, non-response and comparisons with other education data sources to evaluate data quality.

A number of 'Census specific' terms have been used throughout this working paper and the definitions of these are outlined below.

- **dummy forms** - dummy forms are created for dwellings from which Census collectors could not obtain a Census form. Collectors were instructed to record a count of males and females for these dwellings if known. If there was no count, person records were imputed on the basis of data for the rest of the Collector's District. Age, Sex, State of usual residence and Marital status were imputed for persons on dummy forms with all other fields being set to Not Stated or Not Applicable.
- **undercount** - Census counts are subject to inaccuracies because some people have been missed altogether by the Census. Information collected as part of the Post Enumeration Survey (PES) can be compared with information collected in the Census to obtain a measure of the extent to which the Census understated the true population. This measure is referred to in this working paper as undercount.
- **non- response rates** - a non-response rate is normally defined as the proportion of records where a response should have been provided to a given question but is missing. The non-response rate depends on the size of both the 'applicable population' (representing the number of persons who should have answered the question) and the number of Not Stated codes assigned to the question ('question non-response').
- **'topic non-response'** - two of the three education topics in the Census are enumerated using multiple questions; the topic on Attendance at an Educational Institution has two questions, and the topic on Educational Qualifications has five questions. In some cases, Not Stated codes were assigned to all of the questions relating to a given topic. These cases are referred to as 'topic non-response' and they are of particular significance in terms of data quality as it is not possible to make any inference about the status of the person for that topic. In the calculation of non-response rates in 1991, it was assumed that these people did not answer as they saw the questions as irrelevant (that is , did not attend an educational institution or did not have a qualification).
- **2 per cent verbatim file** - a number of questions in the Census require the respondent to write in their response. The 2 per cent verbatim file contains a 2 per cent non random sample of all Census forms and records the response to each write in question answered verbatim. Only records which contain a written response to at least one 'write in response' question can be selected in the sample. These written responses are useful in determining whether people understood a question. This file is used in the analysis of Type of Educational Institution and Qualifications.
- **FURF and IFURF** - there are a number of versions of output files that are produced throughout the various stages of processing of Census data. The Final Unit Record File (FURF) contains Census data in its final format. The Interim Final Unit Record File (IFURF) is an earlier version of the FURF and contains some variables that are useful for analysis of data quality but are not contained on the FURF. In some cases, data

which is referred to on more than one occasion in the working paper may differ if different files were used as the source.

ATTENDANCE AT AN EDUCATIONAL INSTITUTION

Background

All Censuses since the first in 1911 have contained questions aimed at the identification of student numbers. Questions about the type of educational institution being attended by students have also been asked in a number of Censuses to enable more detailed statistics to be produced on the characteristics of people attending educational institutions.

Question Design

In the 1986 Census, a single two-part question (see Appendix 1) was used asking about attendance at an educational institution (Student Status) and type of educational institution attended (Type of Institution). For the 1991 Census this was divided into two separate questions to improve response rates for both questions. The response categories for these questions were essentially unchanged for 1991 except for the combining of 'CAE' with 'University' into a single category with the addition of the words 'or other higher educational institution'.

Other changes made for 1991 were the inclusion of an instruction for school students to mark the second box 'Yes, full-time student' for the question about Student Status and an additional instruction, providing examples of other higher educational institutions, for Type of Institution.

The Pre-school response category for Type of Institution was included in the (Type of Institution) question for the main purpose of improving the quality of Primary school data by ensuring that Pre-schools were not included there. The term 'Pre-school' has different meanings across some States and this makes it difficult to conduct any reliable analysis of Pre-school data. Pre-school data, therefore, will be excluded from the main focus of this working paper.

Census Results

On the 6 August 1991, of the 16,850,334 persons counted in the Census, 4,418,465 were attending educational institutions. A further 607,840 persons gave no response for both Student Status and Type of Institution while the remaining 11,824,029 persons were Not Attending an educational institution as at Census night.

Table 1 provides a summary of the responses for Type of Institution for the 1986 and 1991 Censuses. It shows that 26.2 per cent of the population were attending educational institutions in 1991, an increase of 0.5 percentage points over the 1986 proportion. The increase in student numbers between Censuses, both proportional and nominal, can be attributed to the strong increase in the number of tertiary students. University/CAE students accounted for 3.2 per cent of the population in 1991 (an increase of 1.1 percentage points) while TAFE students accounted for 2.6 per cent in 1991 (an increase of 0.5 percentage points). The proportion of students who recorded 'Other' for Type of Institution also increased between Censuses (0.2 percentage points).

TABLE 1: Summary of responses to Type of Institution, 1986 and 1991 Censuses, all persons, Australia

Response	1986 Census		1991 Census		Difference	
	Count	Per cent	Count	Per cent	Count	Per cent
Type of Institution stated						
Pre-School	279,776	1.8	271,959	1.6	-7,817	-0.2
Infants/Primary	1,564,087	10.0	1,621,313	9.6	57,226	-0.4
Secondary School	1,206,722	7.7	1,176,436	7.0	-30,286	-0.7
TAFE	327,371	2.1	428,884	2.6	101,513	0.5
University/CAE	325,938	2.1	543,930	3.2	217,992	1.1
Other	86,178	0.6	126,524	0.8	40,346	0.2
Total	3,790,072	24.3	4,169,046	24.7	378,974	0.4
Student Status only stated	215,510	1.4	249,419	1.5	33,909	0.1
Total attending*	4,005,582	25.7	4,418,465	26.2	412,883	0.5
Type of Institution & Student Status Not Stated	834,594	5.3	607,840	3.6	-226,754	-1.7
Not Attending	10,762,037	69.0	11,824,029	70.2	1,061,992	1.2
Total	15,602,213	100.0	16,850,334	100.0	1,248,121	0.0

* a person is defined as attending if they were coded as stated Part Time or Full Time for Student Status, or Not Stated for Student Status but stated for Type of Institution

The proportion of student numbers attending other educational institutions decreased between 1986 and 1991. Students attending Infants/Primary schools accounted for 9.6 per cent of the population in 1991 which was a decrease of 0.4 percentage points compared with 1986. Secondary students represented 7.0 per cent of all persons, a decrease of 0.7 percentage points. The proportion of students that did not state Type of Institution was 1.5 per cent which was a slight increase (0.1 percentage points) over 1986.

The significant restructuring that has taken place within the education system in recent years, particularly with respect to tertiary institutions, together with the dynamic nature of the economy and society in general are all factors which could reasonably explain changes in the Census counts from 1986 to 1991. It is also possible that these changes are attributable in part to variations in form design and coding methodology. The following sections will examine the effects that these internal factors have had on 1991 Census data and also look at other education data collections for the purpose of evaluating the quality of 1991 Census data.

Quality of Responses

When comparing 1986 and 1991 Census data on educational institutions, factors such as variation in coding methodologies used and form design can affect changes in data quality. These factors will be examined more closely in this section as well as the incidence of non-response, which is a basis for measuring the quality of Census data.

The shift in the Census date from 30 June in 1986 to 6 August in 1991, while not necessarily affecting data quality, can affect the comparability of the two sets of data, particularly when

comparing age group level data. Where applicable, the shift in Census date will be factored into the analysis to aid comparison in this and later sections.

Editing

Editing of data was conducted during two phases of Census processing.

During the first (main processing) phase, there was some recoding of responses that were inconsistent with each other (eg Not Attending marked but Type of Institution stated) and with other variables, particularly age.

For ages 0 to 3, Type of Institution was recoded to Not Stated for responses other than Pre-school or Other and, for ages 4 to 5, a similar recode was made for responses that did not include Primary, Pre-school or Other.

All persons aged 6 to 13 are within the prescribed ages for which school attendance is compulsory in Australia. For persons aged 6 to 10, Type of Institution was recoded to Not Stated for responses except Primary school or Other. For persons 11 to 13 years old, Type of Institution was recoded to Not Stated for responses except Primary, Secondary or Other. For all persons aged 6 to 13 and coded as Not Attending, Type of Institution was reset to Not Stated and Student Status set to Full Time.

For ages 15 and above, Type of Institution was recoded to Not Stated if it was originally coded Pre-school or Primary. In addition, for ages 20 and above, Type of Institution was reset to Not Stated if it was Secondary.

Further recoding was carried out during the second (output processing) phase. For persons aged 6 to 14, Student Status was recoded to Full Time in cases where student status was originally coded as Not Stated or Not Attending and Type of Institution was stated. For persons aged 0 to 4, Type of Institution was recoded from Not Stated to Pre-school where Student Status was Full or Part Time. A more detailed explanation of coding actions is contained in the summary of the edits performed provided at Appendix 2. Unfortunately there is no information available on the extent of edits performed in 1986.

Non-Response

The adoption of two separate questions for education attendance in the 1991 Census, as already mentioned, was aimed at improving response rates for the two variables, Student Status and Type of Institution. Table 2 shows the breakdown of non-response combinations for the two questions for both 1986 and 1991 Censuses.

TABLE 2: Non-response combinations for attendance at an educational institution, 1991 Census, all persons, Australia

Response	Count		% of Population	
	1986	1991	1986	1991
Student Status & Type of Institution stated	3,457,767	4,145,118	22.2	24.6
Student Status only stated	215,510	249,419	1.4	1.5
Institution type only stated	332,305	23,928	2.1	0.1

Total Attending	4,005,582	4,418,465	25.7	26.2
Topic Not Stated	834,594	607,840	5.3	3.6
Not Attending	10,762,037	11,824,029	69.0	70.2
Total persons	15,602,213	16,850,334	100.0	100.0

The incidence of non-response in 1991 is concentrated most among records that have Not Stated coded to both Student Status and Type of Institution with 607,840 cases or 3.6% of the total population. A disproportionate number of these records are concentrated among persons aged 0 to 4 and this might be because most of this age group is not of school age and people may feel the two questions do not apply to them.

Non-response to Student Status

Student Status was asked of all persons in the Census. Table 3 shows the distribution of 1986 and 1991 Not Stated codes for Student Status by age. Table 3 also shows a comparison of the 1986 and 1991 Census non-response rates for Student Status by age group.

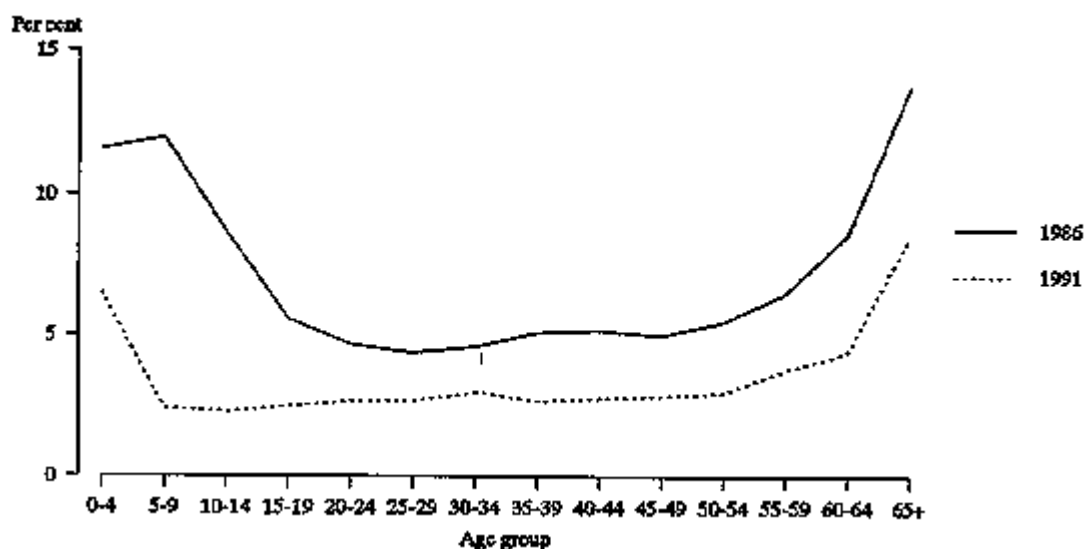
TABLE 3: Non-Response to Student Status by Age, 1986 and 1991 Censuses, all persons, Australia

Age	Count		% of age group	
	1986	1991	1986	1991
0-4	138,274	81,637	11.6	6.5
5-9	139,064	30,300	12.0	2.4
10-14	112,546	28,233	8.7	2.3
15-19	73,579	32,632	5.6	2.5
20-29	117,078	71,851	4.5	2.7
30-39	120,068	76,283	4.9	2.8
40-49	92,415	64,819	5.1	2.9
50-59	87,061	51,688	6.0	3.4
60+	286,814	194,325	12.2	7.4
Total	1,166,899	631,768	7.5	3.7

In 1991, relatively high non-response rates were measured for the 0 to 4 years and 60 years and older age groups, although both were well below the 1986 levels. It is likely that a large portion of the non-respondents within the 0 to 4 and 60 years and older age groups felt that the questions did not apply to them because they were too young or too old to be attending an educational institution. Overall, there was a significant improvement in non-response between the Censuses with the non-response rate falling from 7.5 per cent in 1986 to 3.7

per cent in 1991. This improvement is most evident in the age groups between 5 and 14 but was also significant for the 15 to 19 age group. This is illustrated more clearly in Figure 1 on the following page.

FIGURE 1: Distribution of non-response rates for Student Status by Age, 1986 and 1991 Censuses, Australia



The overall improvement in response could reasonably be attributed to the change in form design for 1991. In 1986, respondents may have been more likely to regard the question about Student Status as a redundant part of what was then a two-part question; particularly if the respondent was a student and was required to give a response to the second part of the question (Type of Institution).

This would explain the high relative non-response rates in 1986 for persons between the ages of 5 and 19. The change from a two-part question to separate questions has probably made the Student Status question seem less redundant to respondents, consequently lowering the incidence of non-response for Student Status between Censuses.

Another element of the form design which was new for the 1991 Census was the instruction for school students to mark the Full-time student box for Student Status. It is possible that this may also have contributed to the level of improvement in the non-response rate for the age groups concerned.

Non-response to Type of Institution

Non-response rate for Type of Institution is normally calculated using the applicable population (the number of persons who should have answered the question based on their responses to Student Status and Type of Institution). This method excludes 'topic non-response' records (who were assumed to not be attending an education institution) and is calculated by dividing the number of respondents who were coded Full time or Part time for Student Status and Not Stated for Type of Institution, by the number of persons who stated Full time or Part time to Student Status plus the number of persons Not Stated to Student Status and stated to Type of Institution.

An alternative way of analysing non-response for Type of Institution is to include 'topic non-response' records and this can be a useful measure for evaluating non-response data for persons of compulsory school age, since it is reasonable to assume that they were at school. For persons above school age level, however, non-response data using 'topic non-

response' records is not comparable because it is more likely, particularly for older persons, that these people are not attending an educational institution. This makes it difficult to evaluate non-response rates across different age groups and, for this reason, non-response rates for Type of Institution which exclude 'topic non-response' records are used in this analysis. (N.B. in the 1986 Census working paper on education, non-response rates for Type of Institution included 'topic non-response' records).

Table 4 shows the distribution of Not Stated codes for Type of Institution by age for 1986 and 1991. The table also shows a comparison of the 1986 and 1991 Census non-response rates for Type of Institution by age group.

TABLE 4: Non-Response to Type of Institution(a) by Age, 1986 and 1991 Censuses, all persons, Australia

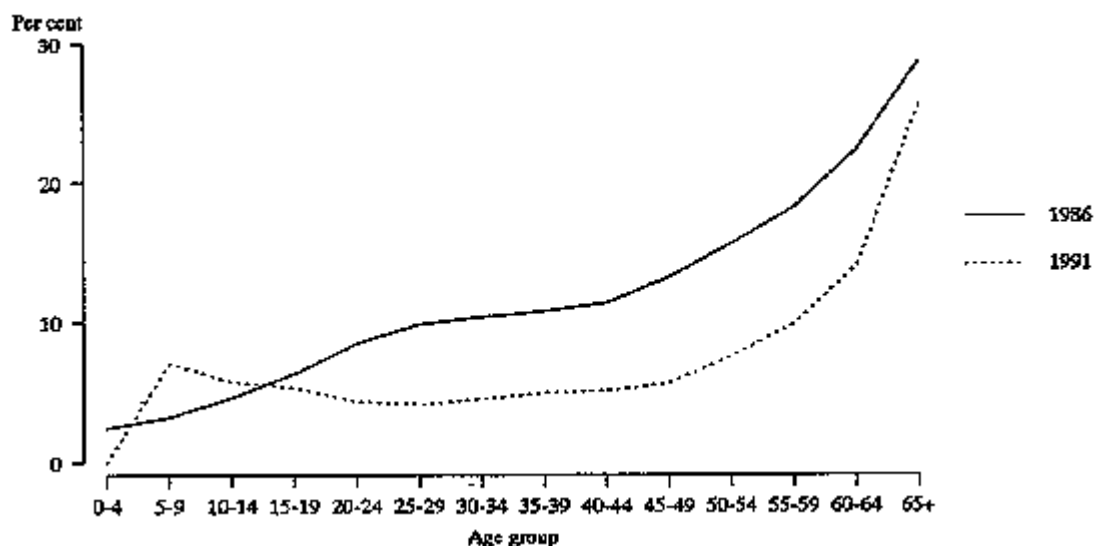
Age	Count		% of students attending	
	1986	1991	1986	1991
0-4	6,005	0(b)	2.5	0.0(b)
5-9	36,620	86,172	3.3	7.1
10-14	59,478	70,238	4.7	5.8
15-19	49,721	48,369	6.4	5.4
20-29	29,508	20,298	9.0	4.3
30-39	17,701	10,955	10.5	4.8
40-49	8,279	6,387	12.1	5.4
50-59	4,138	2,825	16.7	8.4
60+	4,060	4,175	25.7	21.0
Total	215,510	249,419	5.4	5.6

(a) excludes 'topic non-response' records

(b) an edit was made during the output processing phase for persons aged 0 to 4 years to recode Type of Institution from Not Stated to Pre-school where Student Status was stated as attending

The non-response rate for Type of Institution between 1986 and 1991 increased slightly from 5.4 per cent to 5.6 per cent. On an age group basis however, non-response rates have declined for each age group except 5 to 9 and 10 to 14 year olds. Figure 2 shows the distribution of non-response rates for Type of Institution.

FIGURE 2: Distribution of non-response rates* for Type of Institution by Age, 1986 and 1991 Censuses, Australia



* excludes 'double not stated' codes

The increase in non-response for Type of Institution for the 5 to 9 and 10 to 14 age groups, together with the very strong improvement (relative to other age groups) in non-response for Student Status is a pattern that seems to be inconsistent with all other age groups. One possible explanation for this is that the instruction for school students, although improving the response to Student Status for these age groups, may have increased non-response to Type of Institution. Respondents in these age groups may have felt that they did not need to indicate their type of institution as this was implicit both in their compliance with the instruction and response to Student Status.

For statistical purposes it is reasonable to assume, except where age has been imputed, that all 6 to 9 year olds are attending primary school.

Impact of dummy records on topic non-response

Where response was made to Student Status and/or Type of Institution, it is possible to make reasonably accurate assumptions about the student status of the respondent. More detailed analysis is needed, however, to better understand the impact on Census counts of records where no response was given for Student Status and Type of Institution ('topic non-response' records).

Dummy forms are created for dwellings from which Census collectors could not obtain a Census form. Collectors were instructed to record a count of males and females for these dwellings if known. If there was no count, person records were imputed on the basis of data for the rest of the Collector's District. Age, Sex, State of usual residence and Marital status were imputed for persons on dummy forms with all other fields being set to Not Stated or Not applicable.

Table 5 shows the effect of 'dummy forms' (see Section 1) on the number of 'topic non-response' records for the 1986 and 1991 Censuses. In 1991, 3.6 per cent of all records (607,840) were 'topic non-response' records compared with 5.3 per cent (834,594) in 1986. In terms of the overall 'topic non-response' count, therefore, the performance of the education questions in 1991 resulted in an improvement in data quality over 1986. In 1986, approximately 11.9 per cent of 'topic non-response' records (99,348) were persons for whom dummy forms were created. In 1991, a much larger proportion of 'topic non-response' records, 33.3 per cent (202,742), were dummy records. Because of the imputation associated with dummy forms it is not possible to make any reasonable judgement about the student status of these persons. The increase in the number of dummy forms between censuses results in a loss in data quality as there is a greater number of persons for whom

there is no judgement about student status however this outcome is independent of the performance of Census education questions.

TABLE 5: Effect of dummy records on 'topic non-response' records, 1991 Census, Australia Response

Response	Count		% of Population	
	1986	1991	1986	1991
Student Status and/or Type of Institution stated	14,767,619	16,242,494	94.7	96.4
Student Status and Type of Institution Not Stated				
Dummy record	99,348	202,742	0.6	1.2
Other	735,246	405,098	4.7	2.4
Total 'topic non-response' records	834,594	607,840	5.3	3.6
Total Persons	15,602,213	16,850,334	100.0	100.0

For the 1991 Census, there were 405,098 'topic non-response' records in addition to the 202,742 dummy records. Of these 14.0 per cent (56,616) were aged 0 to 4 years and a further 41.4 per cent (167,891) were aged 60 and over. It is reasonable to assume that the majority of these people were Not Attending an educational institution. Another 6.1 per cent (24,651) were aged between the ages of 5 and 14 years and thus could be expected to have been attending school, since it is compulsory in all states for persons within these ages. For the remaining 38.5 per cent (155,940) of persons who had 'topic non-response' records, while it is possible that some were attending an educational institution, a significant number of them may not have responded to either question because they felt that the questions did not apply to them, probably because they were not attending an educational institution.

Non-response rates for both Student Status and Type of Institution by Sex are set out in Appendix 3. Non-response rates for Student Status for 0 to 4 year old male and females are significantly higher than for other age groups. Non-response rises steadily for males and females aged 55 or more. Non-response is slightly higher for males between the ages of 15 and 25. This reflects a general pattern of non-response for males in these age groups which is consistent across a number of Census questions.

Comparison with non-census data

An examination of the discrepancies in attendance counts between the Census and alternative data sources is another means of evaluating the reliability of Census data even though there is some variation in the scope and time frame between the different collections.

There are a number of alternative sources of data on educational institution attendance for each type of institution. These are the ABS Transition from Education to Work (TEW) Survey, National Schools Statistics Collection (NSSC), DEET Higher Education Statistics and the National Council for Vocational Education Research (NCVER) Selected TAFE Statistics. As already mentioned, this analysis will exclude pre-school attendance.

The following analyses use adjustments, where appropriate, to facilitate the comparability of data. Details of how these adjustments were calculated are provided in Appendix 4.

Infant/Primary and Secondary

Data on Infant/Primary and Secondary school attendance are set out in Table 6. For these institutions, Census figures are compared with enrolment data from the 1991 National Schools Statistics Collection (NSSC) published by the ABS (Cat. No. 4221.0). The statistics were compiled from collections of both government and non-government school data. Government school data was collected by State governments who compile the data from their central administrative databases as well as collecting directly from educational establishments. Non-Government school data was collected by DEET who collect directly from establishments. The census date for the NSSC was 2 August 1991 but the reference date for student age was 1 July 1991.

Table 6 shows the comparison of 1991 NSSC student counts for Primary and Secondary schools with the 1991 Census counts. The Census counts of primary and secondary students, for both non-government and government institutions is between 9 and 10 percent lower than the NSSC count. The 1991 Census recorded a number of school age persons Not Stated for Type of Institution and as already mentioned, it can reasonably be assumed that the majority of these persons were attending an educational institution. The Census also missed a number of people on Census night and the Post Enumeration Survey conducted after the Census provides broad-level measures of the extent of this undercount.

TABLE 6: Comparison of 1991 Census and NSSC counts of Primary and Secondary students before and after adjustments for Census undercount and Not Stated codes, all persons, Australia

	Govt	Primary Non-govt	all schools	Govt	Secondary Non-govt	all schools
NSSC	1,338,533	447,913	1,786,446	878,693	409,998	1,288,691
Census count	1,219,189	402,124	1,621,313	795,402	381,034	1,176,436
Difference before adjustments	-119,344	-45,789	-165,133	-83,291	-28,964	-112,255
Percentage	-8.9	-10.2	-9.2	-9.5	-7.1	-8.7
Adjustment for undercount*	17,555	5,771	23,326	14,522	6,965	21,487
Adjustment for Not Stated*	122,960	40,415	163,375	65,644	31,370	97,014
Difference after adjustments	21,171	397	21,568	-3,125	9,373	6,248
Percentage	1.6	0.1	1.2	-0.4	2.3	0.5

* The methodology used in the calculation of these adjustments is detailed in Appendix 4.

Methodological differences between the data collections are responsible for substantial differences between the two data sets. The two main factors are the existence of both

undercounting and Not Stated codes in the Census. To demonstrate their potential effect, pro-rata allocations of Census undercount and Not Stated codes have been calculated and included in Table 6 (see Appendix 4 for details about these adjustments). As can be seen in the table a significant amount of the variation between the two collections could be explained by these two factors.

Comparative analysis by age groups was investigated but the results were difficult to interpret because of the different age reference dates.

TAFE

The main potential sources for TAFE data are the Selected TAFE statistics produced by NCVER and the Transition From Education To Work survey produced by the ABS (No. 6227.0). The conceptual differences between the Census and NCVER collections are so great as to make comparison of the data meaningless and therefore no comparative analysis of the Census and NCVER data will be done in this paper.

The ABS Transition from Education to Work (TEW) survey is an annual survey conducted in association with the ABS monthly labour force survey. The TEW surveyed persons aged 15 to 64 about their student status and the type of educational institution they were currently attending or had attended within the previous year.

In 1991, the TEW survey was run in May, around three months before the Census. Apart from age, there are a number of other minor differences in scope between the TEW survey and the Census. Overseas residents in Australia and members of Australia's permanent defence forces were not included in the scope of the TEW survey but were included in the Census. However it is likely that the number of persons within these groups who were attending TAFE is not significant and so would not substantially affect the difference between the two collections.

Table 7, shows the comparison between estimates of TAFE students from the TEW survey and the Census count of TAFE students. The table shows the unadjusted Census count and an adjusted Census figure after allowing for Census undercount and Not Stated codes.

TABLE 7: Comparison of 1991 Census count and TEW Survey estimates of TAFE students ages 15-64 years, Australia

	Unadjusted	Undercount	Adjustments* Not Stated	After Adjustments
TEW	518,885	-	-	518,885
Census	425,062	11,383	31,215	467,660
Difference	-93,823	-	-	-51,225
Percentage	-18.1	-	-	-9.9

* The methodology used in the calculation of these adjustments is detailed in Appendix 4.

The unadjusted Census count for all TAFE students aged 15 to 64 is 18.1 per cent or 93,823 persons less than the TEW estimate. After adjusting the Census count for undercount and Not Stated the variation between the counts is almost halved with the Census count 9.9 per cent or 51,225 persons less than the TEW estimate.

It is likely that a number of TAFE students were coded as 'Other' for Type of Institution and this would explain part of the remaining variation between Census and survey data. In 1991 there were 126,524 records with 'Other' coded for Type of Institution compared with 86,178 in 1986; this represented an increase of 31.9 per cent. An examination of the '2 per cent verbatim file' (as defined in Section 1) identified a large amount of responses that were TAFE institutions. Of all records in the file where 'Other' had been coded and a verbatim response had been recorded, around 12 per cent of responses were Australian TAFE institutions.

It is also possible that part of the variation in TAFE student numbers between the Census and survey is due to respondents incorrectly marking University/CAE because they were doing a TAFE administered course held at a University/CAE campus or because of ambiguity associated with the name or type of institution.

Seasonal variation in TAFE enrolments between May and August and sampling error associated with the survey estimate would also be factors contributing to the variation between numbers of TAFE students in the two collections.

Universities/CAE's

Census data on University and CAE attendance has been compared with both the DEET Selected Higher Education Statistics, 1991 and the 1991 Transition From Education To Work survey data produced by the ABS (No. 6227.0).

The DEET Selected Higher Education Statistics are compiled from the annual statistical collections of the Department of Employment, Education and Training. Differences in methodology are likely to be major contributors to any variation between the two collections. DEET data are compiled from administrative records kept by the relevant educational institutions whereas Census data is generated by the responses of people to questions in the Census about education. This creates the potential for differences to arise because some people who classify themselves as University/CAE students in the Census may not be included in the DEET statistics.

DEET data relates to the number of student enrolments and is collected on the basis of the primary enrolment of a student to avoid counting a student more than once. The reference date for the DEET collection for number of students enrolled was 31 March 1991. The age analysis of DEET enrolments data, however, uses the student's age as at 31 December 1990.

Table 8 shows a comparison of Census and DEET counts for University/CAE students. The table compares the DEET count with both the unadjusted Census count, and the estimated Census figure after adjusting for undercount and Not Stated codes. The adjustments were calculated using the methodology outlined in Appendix 4. Overall, the unadjusted Census figure is fairly close, at only 1.8 per cent or 9,392 persons more than the DEET count of University/CAE students, for all ages.

TABLE 8: Comparison of 1991 Census and DEET Selected Higher Education Statistics counts of University/CAE students by age group, all persons, Australia

	Unadjusted	Undercount	Adjustments* Not Stated	After Adjustments
DEET	534,538	-	-	534,538

Census	543,930	15,616	39,217	598,763
Difference	9,392	-	-	64,225
Percentage	1.8	-	-	12.0

* The methodology used in the calculation of these adjustments is detailed in Appendix 4.

After adjustments for Census undercount and Not Stated codes are considered, the difference between the two counts increases considerably with the adjusted Census figure becoming 12.0 per cent greater, overall, than the DEET count.

This percentage represents a difference of around 64,000 students. Although it is not possible to identify any one main factor which would account for this difference, there are a number of possible factors that may at least partly explain this difference.

To begin with, the implicit assumptions underlying the Census adjustments (see Appendix 4) may lead to inaccurate adjustments. The adjustments may overestimate the number of University/CAE students amongst the people not counted in the Census and/or overestimate the proportion of those not answering this question who are University/CAE students.

Also, as previously mentioned, variation may be due to the differences in methodologies. As a result, persons may have been counted in the Census but not included in the DEET data for a number of reasons.

Firstly, respondents may have been confused about the specific tertiary classification of their course, particularly in light of the recent changes that have occurred within this section of the education system. For example, in NSW and Queensland there were TAFE colleges offering advanced education courses, of which, most were excluded from the DEET count but it is possible that a number of these students were coded as University/CAE in the Census.

Also, students may be enrolled in courses administered by universities/CAE's which in an academic sense are not university/CAE courses and so fall outside the scope of the DEET collection. The Royal Melbourne Institute of Technology (RMIT), for example, is primarily a University but it does have a sector within it that offers TAFE courses.

As mentioned earlier, the DEET collection uses 31 December 1990 as the reference date for age data compared with 6 August 1991 for the Census. This difference prevents any meaningful comparative analysis between the two collections at an age group level. The different dates lead to different age group level data between the two collections because there are some students that can be classified in different age groups depending on which date is used. To illustrate the potential impact of the differences in age reference dates on the comparability of the two sets of data, pro-rata adjustments have been made and included in Appendix 5.

Table 9 shows the comparison of Census count and TEW estimate of students attending Universities/CAEs. The table compares the TEW estimate with both the unadjusted Census count and the estimated Census figure after adjusting for undercount and Not Stated codes. Adjustments were made using the same methodology as outlined in Appendix 4. The unadjusted Census count is 8.0 per cent less than the survey estimate, but the table shows that undercount and non-response are factors that could account for much of the variation between Census and survey data. A comparative analysis by age group which illustrates the potential impact of the different age reference dates on the two data sets at age group level

is provided in Appendix 5.

TABLE 9: Comparison of 1991 Census count and TEW Survey estimates of University/CAE student numbers, persons aged between 15 and 64, Australia

	Unadjusted	Undercount	Adjustments* Not Stated	After Adjustments
TEW	588,076	-	-	588,076
Census	541,051	15,575	38,130	594,756
Difference	-47,025	-	-	6,680
Percentage	-8.0	-	-	1.1

* The methodology used in the calculation of these adjustments is detailed in Appendix 4.

Other potential sources of variation between the Census and survey data are: seasonal factors affecting University/CAE enrolments in the period between the two collections, University/CAE students who were overseas residents in Australia, other members of the population that were out of scope for the TEW survey, and sampling error associated with the TEW survey.

There is the potential for further variation, in addition to that described so far in this comparative analysis of University/CAE student data collections, to arise from the number of University/CAE students who might have given incorrect responses in the Census to Type of Institution, such as those who were coded as 'Other' or Not Attending. It is not possible to quantify this error, although it is possible to gain some idea of the incidence of responses coded as 'Other' by examining the 2 per cent verbatim file.

Of all records in the file, records that had been coded 'Other' for Type of Institution and an associated verbatim response had been recorded, around 27 per cent of responses were Australian University/CAE institutions. This number of responses represents about 10 per cent of all records in the 2 per cent verbatim file coded as 'Other' for Type of Institution (with and without associated verbatim responses).

Of the three data sources used in this analysis of University/CAE data, the Census and TEW data are derived using similar collection methodologies, but there are some conceptual differences between the Census and DEET Higher Education Statistics collection affecting the comparability of these two data sets. Firstly, the DEET data is compiled from records of students enrolled in designated University/CAE courses, whereas the Census data is based on the respondent's assessment of whether they are University/CAE students. The difference in data collection methods used is probably creating differences in scope between the Census and DEET data.

In summary, it is likely that the quality of Census data on University/CAE students is reasonably good allowing for the uncertain effects of undercount and non-response. After making allowance for these factors, Census data is very similar to the TEW estimates.

Summary

There was a big improvement in response to Student Status between the 1986 and 1991 Censuses, with non-response falling from 7.5 per cent to 3.8 per cent. This is probably mainly attributable to the separation of the single two-part question into separate questions.

The non-response rate for Type of Institution at 5.6 per cent, while not exceptionally high, was slightly higher than the rate in 1986 of 5.4 per cent. The inclusion of the instruction to students could have been a factor that contributed to non-response for Type of Institution in 1991. Nevertheless, it would seem that the changes made to the question design for the 1991 Census, have resulted in an overall improvement in the quality of educational institution data.

The extent to which non-response has affected the quality of 1991 Census counts of educational institution attendance can be assessed by comparing Census data with other data sources on educational institutions. The comparisons used in this analysis suggest that Census non-response may have contributed significantly to the differences between Census data and the other sources.

Comparison between the National Schools Statistics Collection data and Census data on Primary and Secondary students showed that Census data is relatively understated. The variation between the two collections, however, was reduced substantially when Census undercount and Not Stated codes are taken into account.

The ABS Transition From Education To Work (TEW) survey data on TAFE students was used to evaluate Census TAFE data. There was significant variation between the unadjusted Census count and the TEW estimate. Nearly 50% of this variation, however, may be attributable to Census undercount and Not Stated codes.

Two alternative data sources were used to evaluate Census University/CAE student counts. The Department of Employment, Education and Training's (DEET) Higher Education Statistics count of students is very closely matched with the unadjusted Census count. After making allowance for Census undercount and Not Stated codes, however, there is significant variation between Census and DEET counts. The TEW estimate of University/CAE student numbers was very close to the Census count but only after allowing for undercount and Not Stated codes. Conceptual and methodological differences between the three collections contribute to differences in results and make data quality judgements about Census counts difficult. Also, the considerable restructuring of the education sector which took place between Censuses may have increased confusion amongst respondents about the classification of courses and institution types.

AGE LEFT SCHOOL

Background

A question on Age Left School has been included in the Census since 1976. In the two Censuses prior to that, a question had been asked on level of schooling. Although instructions varied between States and Territories to accommodate the different education systems within a consistent framework, there was a considerable inconsistency in responses. As a result a question on Age Left School has been used as a proxy for level of schooling since the 1976 Census.

Question Design

The question design in 1991 was considerably different to that used in previous Censuses. With the introduction of Optical Mark Recognition (OMR) in 1991, respondents were required to mark the appropriate response category rather than write in their age left school. The impact of the introduction of self-coding will be discussed further in later sections.

Little assistance was provided in the hotline guide to help people interpret this question. People who had a break in their schooling would probably have been advised to enter the

age that they first left school. (In line with the new ABS standard for this variable, this will change in 1996 when an instruction will be included on the form telling people to enter the age they last left school.)

The data was captured by the OMR readers and not subjected to any checks or edits, other than to resolve multiple or missing marks and to ensure that Age Left School was 'Not Applicable' for persons aged under 15 years.

Census results

Many factors may influence the comparison of responses to Age Left School between Censuses. These include factors affecting the population such as death and migration, changing trends in educational attainment and changes in form design. Table 10 below compares responses to the Age Left School question in the last three Censuses.

TABLE 10: Age Left School, Persons aged 15 years and over, Australia, 1981, 1986 and 1991 Censuses

	1981 Census No.	%	1986 Census No.	%	1991 Census No.	%
Still at school	462,487	4.2	563,444	4.7	618,806	4.7
Did not go to school	88,267	0.8	96,943	0.8	122,770	0.9
14 years or younger	2,674,119	24.5	2,444,783	20.4	1,964,839	15.0
15 years	2,690,753	24.6	2,763,197	23.1	2,738,027	20.9
16 years	2,117,625	19.4	2,387,038	19.9	2,508,076	19.2
17 years	1,400,171	12.8	1,730,536	14.5	2,114,757	16.2
18 years	755,432	6.9	971,546	8.1	1,218,854	9.3
19 years or older	247,325	2.3	351,948	2.9	758,811	5.8
Not stated	483,245	4.4	655,875	5.5	1,040,712	8.0
Total	10,919,424	100.0	11,965,310	100.0	13,085,652	100.0

Between 1981 and 1991, particularly between 1986 and 1991, there were substantial increases in the proportion of people responding '19 years or older' and in the number of people not stating their Age Left School. Responses of younger ages, in particular '14 years or younger' declined. Increases in the proportion of people giving an age of 17 years or older for Age Left School probably in part reflect the increasing school retention rates over the period. However, as mentioned above, other factors including form design may also be involved.

Quality of Responses

Non-response rates

There was an increase in the non-response rate to Age Left School from 5.5 per cent in the 1986 Census to 8.0 per cent in the 1991 Census. This substantial increase indicates that more people had difficulty in responding to this question in 1991 than in 1986. Table 11 below shows that an increase in the non-response rate occurred in nearly all age groups.

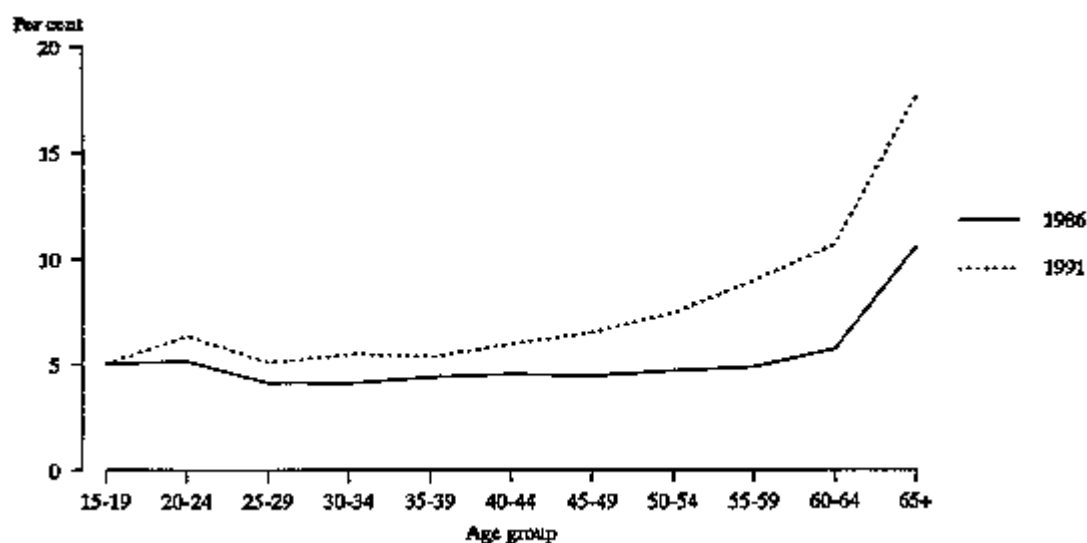
TABLE 11: Non-response to Age Left School by Age, persons aged 15 years and over,

1986 and 1991 Censuses

	1986 Census			1991 Census		
	Not Stated codes	Total persons	Non-resp. rate	Not Stated codes	Total persons	Non-resp. rate
15-19	66,337	1,317,266	5.0	66,327	1,314,181	5.0
20-24	66,293	1,281,664	5.2	85,278	1,343,820	6.3
25-29	53,889	1,297,967	4.2	95,265	1,336,017	5.1
30-34	50,827	1,233,570	4.1	76,025	1,387,223	5.5
35-39	54,168	1,234,265	4.4	68,944	1,293,366	5.3
40-44	44,923	986,870	4.6	75,270	1,261,380	6.0
45-49	36,558	818,916	4.5	65,149	1,001,396	6.5
50-54	33,497	812,727	4.7	60,949	821,346	7.4
55-59	35,875	735,972	4.9	63,345	705,237	9.0
60-64	39,887	699,386	5.7	76,264	714,954	10.7
65+	173,625	1,646,718	10.5	334,896	1,906,732	17.6
All ages	655,879	11,965,321	5.5	1,040,712	13,085,652	8.0

The magnitude of the increase is shown even more clearly when the data is presented in a graph. In particular, the non-response rates for older people, although slightly higher in 1986, become much higher in 1991. Whereas in 1986, 38.0 per cent of Not Stated codes were for people aged 55 years and older, in 1991 the percentage rose to 45.6.

FIGURE 3: Non-response rates to Age Left School, by age, 1986 and 1991 Censuses



The increase in non-response rates occurred in all age groups except 15 to 19 years, indicating that generally people found this question more difficult to answer in 1991 than in 1986. The rate for people in older age groups (60 years and older) in both Censuses was much higher than other age groups, a pattern also observed for other questions, such as those concerning labour force participation. This may be because older people do not see such questions as being relevant to them. An additional problem with the Age Left School question may be the difficulty of recalling their exact age left school. The increase in the non-response rates for Age Left School between 1986 and 1991 could indicate that the new form design may have presented older people with greater difficulties.

The non-response rate for 20 to 24 year olds is higher than the surrounding age groups,

particularly in 1991, and this could indicate specific problems for this age group, particularly as these people would have recently left school and thus should be able to respond to the question with little difficulty. Analysis of non-response to Age Left School by Type of Educational Institution attended for this age group showed a high level of non-response to Age Left School among University/CAE students: 9.2 per cent compared to 5.9 per cent for the remainder of the population aged 20 to 24 years. Thus, it is possible that students attending tertiary institutions may have not been sure how to answer this question, particularly if they understood 'school' to include some tertiary institutions, and so did not respond.

The overall increase in the non-response rate is of concern and indicates that people had problems with the question design used in 1991. The introduction of self-coding was not shown to have a detrimental effect on other questions in the Census, although non-responses rates for many questions did increase, so this alone would not seem to be the reason for the high non-response. There are other changes however which may possibly have affected the non-response rates.

Firstly, the response categories did not clearly indicate what the question was asking, unlike 1986 where 'Age left school' was written at the top of the box. Thus, if people did not read the question on the left hand side, the response categories would not have been clear and as a result some people may have skipped the question. Secondly, the question on Age Left School was preceded by the question 'Is the age given for the person 15 years or more?'. It was noted in observation studies that this question confused some people as age had already been asked earlier on the form. The confusion may have made them more likely to skip the next question on Age Left School if it appeared at all unclear.

Inconsistent response patterns

One of the response categories for Question 23 on Age Left School is 'Still at primary or secondary school'. Information on whether a person is attending secondary school is also collected as a part of the two other groups of questions on education in the Census: Question 21 concerning attendance at an educational institution and Question 24 concerning qualifications obtained. Comparison with these responses reveals that there was some confusion concerning Age Left School. In particular, some people who had said they were attending a secondary school in Question 21 gave an age left school for Question 23.

The tables below compare Age Left School responses of people who made this error with their response to Q24 and their actual age. Over 99 per cent of people with this error were aged under 20, hence persons 20 years and older have been excluded from the tables. Of all people aged 15 to 19 years, 1.1 per cent made this error. Data from the Interim Final Unit Record File (IFURF) was used for the comparison with Question 24 as this information was not stored on the Final Unit Record File (FURF) which was used as the basis for most of the analysis (see Section 1 for definitions of IFURF and FURF).

TABLE 12: Responses to Question 24 on Educational Qualifications, persons with response of 'Secondary school' for Q22 but with an age stated for Age Left School, aged 15 to 19 years

Question 24	Actual age					Total
	15	16	17	18	19	
No	890	1,187	1,500	1,319	1,023	5,919

Still at primary or secondary school	1,284	1,397	1,576	720	242	5,219
Still studying for first qualification	115	243	374	435	325	1,492
Yes	52	59	95	183	234	623
Not Stated	528	438	353	192	108	1,619
Total	2,869	3,324	3,898	2,849	1,932	14,872

Source: IFURF

Over a third of people in the above table responded 'Still at primary or secondary school' for Question 24 on Educational Qualifications. Thus, these responses supported their responses to the question on Educational Institution, indicating that Age Left School responses were incorrect. Many of the people responding 'No' to Question 24 may also have been still at secondary school. Table 13 below compares actual age with Age Left School.

TABLE 13: Responses to Question 23 on Age Left School, persons with response of 'Secondary school' for Q22 but with an age stated for Age Left School, aged 15 to 19 years

	Actual age					
Question 23	15	16	17	18	19	Total
14 or younger	156	94	98	63	39	450
15 years	2,122	359	341	230	189	3,241
16 years	246	2,293	493	377	353	3,762
17 years	140	310	2,479	721	496	4,146
18 years	82	116	329	1,259	416	2,202
19 or older	81	107	133	185	439	945
Total	2,827	3,279	3,873	2,835	1,932	14,746

Source: FURF (SUPERCROSS)

For over 55 per cent of people in this table, the stated age left school is the same as their actual age. This suggests that many may have misunderstood the question. One possible reason for this may involve the question immediately before the Age Left School question, which asked age for the second time on the form and appeared to confuse some people.

Other errors are more difficult to explain. For example, over 10 per cent of people in Table 13, excluding persons who had age imputed, gave an age left school greater than their actual age. It is not clear how these people interpreted the question, although they may have given the age they intended to leave school. It is also possible that some of the inconsistencies resulted from errors made in responding to the questions on age or educational institution.

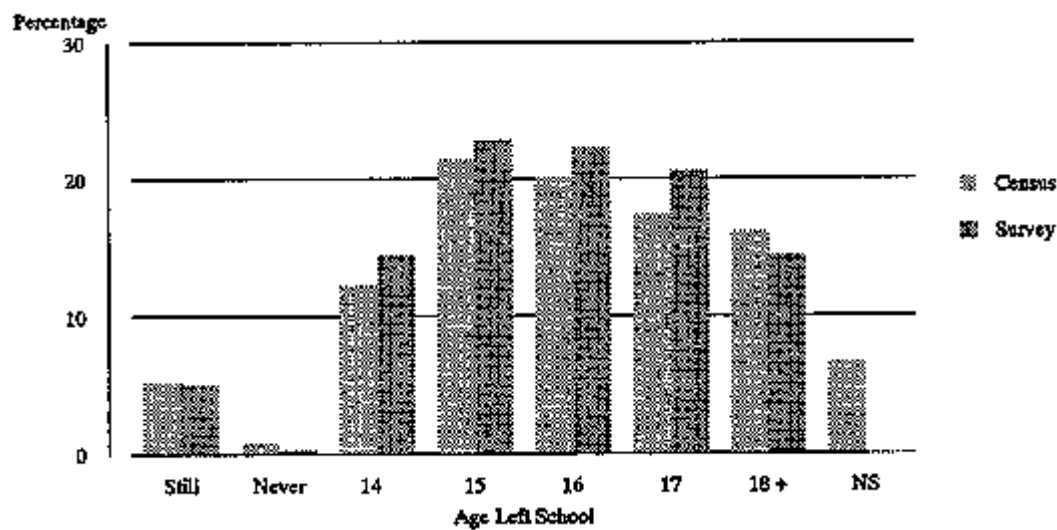
Although the impact of these apparent errors on the overall counts is slight, they are of concern because they indicate that some people have had difficulties answering these

questions and perhaps understanding the concept of Age Left School. There may also be other instances in other age groups of misreporting to one or other of the three questions examined but unfortunately these can not be identified.

Comparison with Educational Attainment Survey data

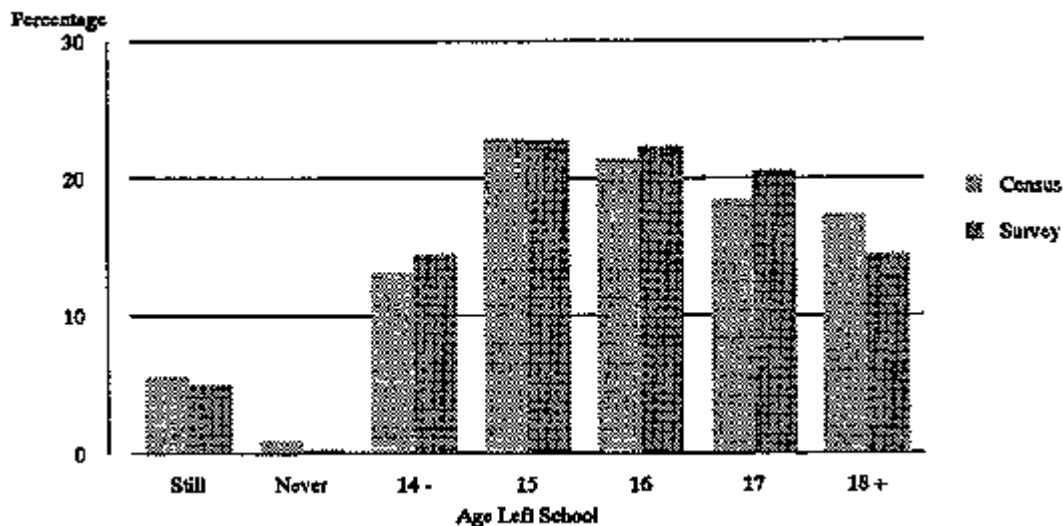
Data on Age Left School was collected as a part of the Labour Force Status and Educational Attainment Survey which was conducted in February 1991. This survey collects data on the highest educational qualifications of persons aged between 15 and 69 years. As a part of this Survey, information was obtained on the age left school for all persons included in the Survey. Figure 4 below compares the data for the Survey with Census data. The Census categories of '18 years' and '19 years or older' were combined in the Survey.

FIGURE 4: Percentage distribution of responses to Age Left School, February 1991 Educational Attainment survey and August 1991 Census, persons aged between 15 and 69 years



The impact of the high non-response rate is evident in this graph. The magnitude of non-response is such that it would be a large factor in the differences between Census and Survey proportions in the categories of 17 years and younger. The Census Not Stated category is excluded in the graph below to enable a comparison of stated responses.

FIGURE 5: Percentage distribution of responses to Age Left School, excluding Not Stated codes in the Census, February 1991 Educational Attainment survey and August 1991 Census, persons aged between 15 and 69 years



The largest difference now is for the category '18 years or older' where the Census count is 10 per cent higher than the Survey estimate. This is a cause for concern as the Census count would be expected to be lower because it is still subject to undercount, that is, the small proportion of people who are not included on a Census form.

The Census count of persons leaving school aged 18 years or older increased by 49.4 per cent between 1986 and 1991, principally due to the count of persons leaving aged 19 years or older which doubled. The size of this increase together with the anomaly shown for the category in Figure 5 suggest the 1991 Census Age left School data may contain deficiencies. Further evidence of this, together with a possible explanation, is contained in the cohort analysis of Section 2.4.

The difference in methodology between the Survey, which is conducted by an interview, and the Census, which is self-enumerated, may in part explain the differences between the data. In the Survey, the interviewer could clarify the definition of 'school' if required while in the Census little assistance was available. Thus, in the Census some people may have mistakenly included tertiary institutions as 'school', particularly as the term 'school' has different meanings throughout Australia and in other countries.

Another difference between the Census and Survey data evident in Figure 5 is the slightly greater proportion of people in the Census indicating that they did not attend school. It is possible that some people may have misinterpreted the question as referring only to their present situation. Alternatively, some people may have felt more sensitive about revealing that they did not attend school when talking to an interviewer than when completing a Census form themselves. This category is still very small however and does not have much impact on the overall distribution.

Comparison with data for Level of Schooling

Apparent Retention Rates

As mentioned in Section 2.1.1, there has been an increase in the proportion of people giving age left school as 17 years or older between the 1986 and 1991 Censuses. This is at least partly due to changes in society. The trend in the length of time people stay at school can be assessed using data on Apparent Retention Rates. These rates were obtained from the National Schools Statistics Collection (NSSC) and represent the percentage of students who continued to Years 10, 11 or 12 in a particular year from their respective cohort groups at the

commencement of their secondary schooling (Schools Australia 1991, ABS Cat. No. 4221.0). Table 14 shows a steady increase in Apparent Retention Rates at Years 10, 11 and 12 between 1981 and 1991.

TABLE 14: Apparent Retention Rates a for Australia, 1981, 1986 and 1991

	1981	1986	1991
	%	%	%
Year 10	91.4	94.1	98.8
Year 11	55.2	68.3	86.0
Year 12	34.8	48.7	71.3

a Proportion of a cohort group commencing secondary school that have continued to the level shown in a particular year.

Source: 'Schools Australia', ABS Cat. Nos. 4202.0, 4221.0

Thus, the increased proportion of responses to Age Left School of 17 years and older reflects, to some extent, an increase in the time people spend at school. It is not, however, possible to directly quantify the effect of the increase in retention rates on Age Left School. There is a lag as the Apparent Retention Rates refer to current students while Age Left School refers only to people who have completed their schooling. There are also difficulties comparing level and age data, as is discussed further below.

Correspondence between level of schooling and Age

Age Left School is used as a proxy for level of schooling attained and, while age and level of schooling are related, the adequacy of the proxy is affected by a number of factors. The education systems vary between States and Territories, affecting the average age at which students complete each 'level'. The distribution of ages in any level covers several years due to factors including students skipping or repeating years or moving between different systems.

The table below shows the effect of the differences between groups of States and Territories in Australia. The first group of States and the ACT have similar education systems while, in the second group, Queensland and Western Australia students spend one year less in secondary school and South Australia and the Northern Territory have forms of continuous enrolment which affect the age distribution of students. The data has been adjusted to represent estimated age at the end of the year.

TABLE 15: Percentage distribution of Age by Level of School, students aged 14 years or older, National Schools Statistical Collection 1991

	Approximate age at end of year *						
	14 years	15 years	16 years	17 years	18 years	19 years or older	All students 14 or older
Year 10							
NSW, Vic, Tas, ACT	1.0	39.0	47.9	10.7	0.9	0.6	100.0
Qld, SA, WA, NT	19.8	46.0	29.5	3.6	0.5	0.6	100.0

Year 11							
NSW, Vic, Tas, ACT	0.0	1.1	37.9	46.7	11.3	3.0	100.0
Qld, SA, WA, NT	0.1	19.8	44.9	28.8	4.3	2.1	100.0
Year 12							
NSW, Vic, Tas, ACT	0.0	0.0	1.2	37.7	46.3	14.8	100.0
Qld, SA, WA, NT	0.0	0.1	18.9	42.4	28.3	10.3	100.0

* The reference date for age in the NSSC was 1 July. The data have been adjusted the represent age at the end of the year by assuming that half the students would have a birthday between July and the end of the year.

While the majority of students in Year 12 in the first group are 18 years or older, the majority in the second group are 17 years or younger. Thus, when comparing age left school between States and Territories, both differences in schooling systems and differences in retention rates will be reflected. The table also shows that the ages of students in any level are spread across a three or four year span and so a student leaving at age 17 may have completed Year 10, Year 11 or Year 12. This variation is less extreme for young and old age groups, with almost no people aged 15 having completed Year 12, while most people aged 19 years or older have completed Year 11 or higher.

While Age Left School may provide the best possible data on level of schooling that can be collected in the Census, when analysing the data, in particular comparing results for States and Territories, it has limitations which must be taken into account.

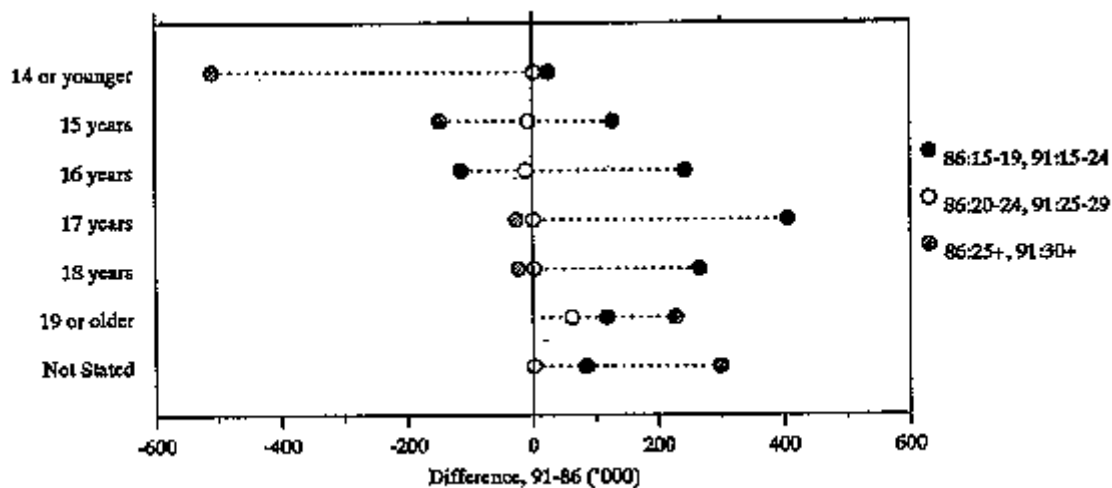
Cohort analysis, 1986 and 1991 Censuses

All persons 15 years and older

It is possible there are factors which affected the comparison of counts between 1986 and 1991 in addition to the increase in retention rates. To explore these further, an analysis by age cohort groups will be used. Cohort groups are groups which should contain the same people between Censuses, for example, persons aged 30 to 34 years in the 1986 Census would be aged 35 to 39 years in the 1991 Census.

Figure 6 below shows the difference between 1986 and 1991 counts for Age Left School responses for three groups. The smallest categories of 'Still at primary or secondary school' and 'Did not go to school' are not included in this analysis. The first group is not strictly a cohort group as it contains persons aged 15 to 19 in 1986 and 15 to 24 in 1991. Changes for this group between 1986 and 1991 reflect the impact of school leavers in this period. Most of the people in the older cohort groups should have completed school before the 1986 Census.

FIGURE 6: Difference between 1986 and 1991 Census counts for Age Left School, by Age.



As most of the difference between the 1986 and 1991 for the categories for 15 to 19 years is due to the youngest group, this change mostly reflects the impact of people completing their schooling.

This is not the case for the response category of '19 years or older', however, where the other age groups also had a large impact. Most people in the age cohort aged 25 to 29 years in 1991 would have completed their schooling by 1986 so the increase in people in this cohort having left school aged 19 years or older appears very large. The proportion of this cohort having an Age Left School of 19 years or older increased from 3.7 per cent in 1986 to 8.2 per cent in 1991 which suggests that some people may have answered this question incorrectly in 1991, possibly including tertiary institutions as 'school'.

The oldest cohort consists of persons aged 25 years or older in 1986 and 30 years and older in 1991. The size of this cohort group has fallen between the Censuses due to deaths although this is not the case for all categories. While the decrease in responses of 14 years or younger seems very large, there is a large increase in responses of '19 years or older' and in Not Stated codes. Most people in the oldest cohort group should have completed their schooling before 1986 and therefore the increase in responses of '19 years or older' is probably due to other factors. A more detailed cohort analysis for this group is included in the next section to further examine the large changes observed.

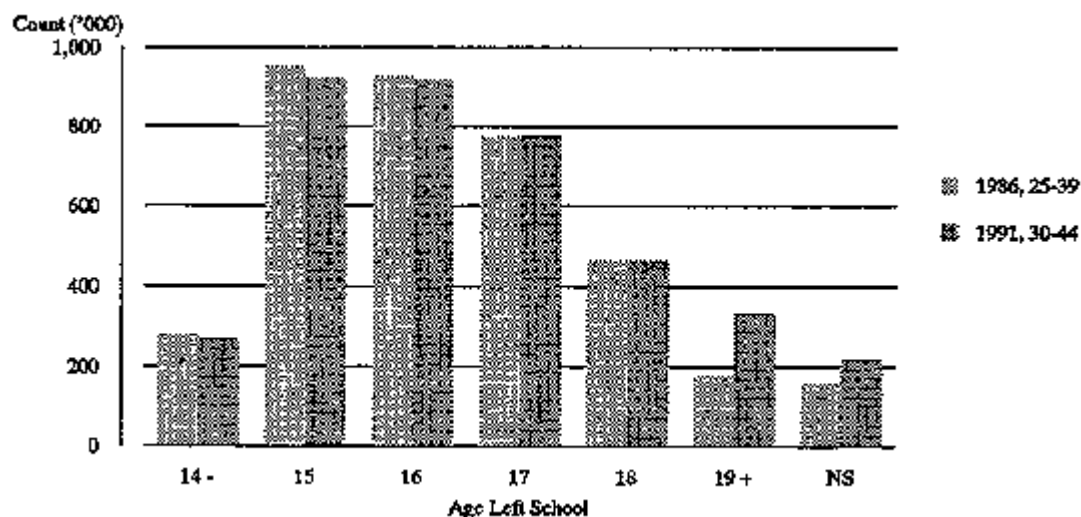
Cohort analysis for persons 25 years and older in the 1986 Census, 30 years and older in the 1991 Census

In this analysis, 1986 and 1991 counts of responses to Age Left School will be compared for three smaller cohort groups of people: those aged 25 to 39 years in 1986 and 30 to 44 years in 1991, those aged 40 to 54 years in 1986 and 45 to 59 years in 1991 and those aged 55 years and older in 1986 and 60 years and older in 1991.

The graphs below show the counts for each of the three cohorts in 1986 and 1991. Differences between the cohorts due to long-term changes in retention rates are apparent in the differences between the graphs, with the youngest cohort having higher ages at leaving school than the others, particularly the oldest group where many people left school aged under 14 years.

Figure 7 below shows the 1986 and 1991 Census counts for the youngest cohort in this analysis.

FIGURE 7: Age Left School counts, persons 25 to 39 in the 1986 Census and persons 30 to 44 in the 1991 Censuses

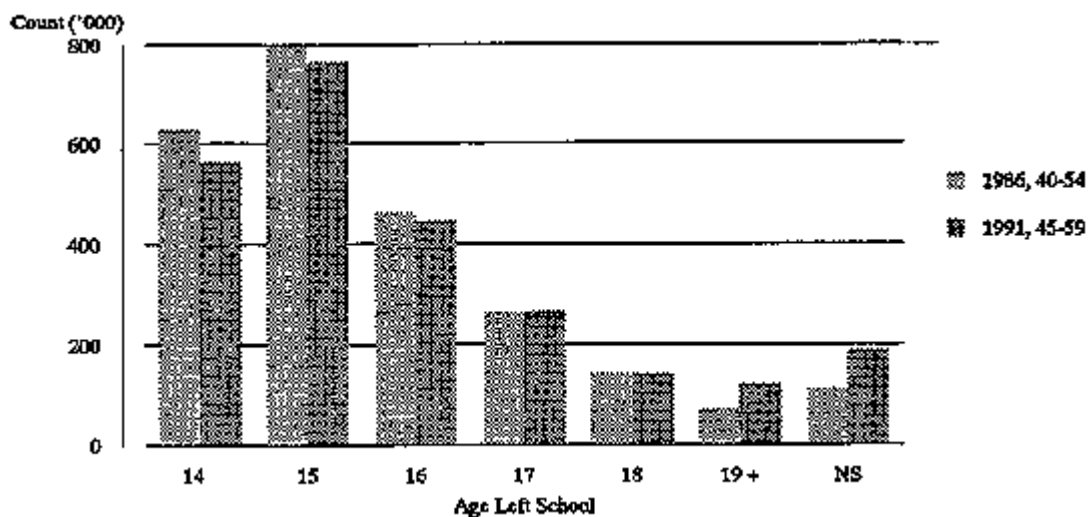


The counts for most response categories for the youngest cohort are very close with the exception being the '19 years or older' category, which increased by 156,447 people. In part, this could be due to people immigrating from overseas, however only 80,822 people in this cohort group who arrived in Australia between 1986 and 1991 gave an age left school of '19 years or older'. Thus, the increase is also apparently due to people responding differently in 1991 than in 1986. Some of the change could be due to people returning to complete their schooling between the Censuses, but the numbers involved seem too large for this to be the sole explanation.

Some of the possible causes of the increase in response to '19 years or older' might be found in form design changes. As discussed earlier in Section 2.2.1, in 1991 the question for Age Left School followed a question asking whether the person was younger or older than 15 which may have caused some confusion. In addition to possibly increasing non-response to the succeeding question, this may also have resulted in some people marking their current age ('19 years or older') rather than the age they left school. A similar tendency was observed in Section 2.2.2 for persons aged 15 to 19 years.

The most significant change between 1986 and 1991 was the introduction of OMR technology for data capture, which required people to mark boxes rather than write in answers. The different format could have several results which were previously investigated in 'Census Working paper 93/2, Comparison of self-coded and write-in responses: July 1992 Test'. The presence of a list of responses may have prompted people to give certain responses. The term 'school' was not defined and an impression that post-secondary studies were meant to be included may have been reinforced by the inclusion of the '19 years or older' category, since most people leave secondary school younger than 19 years of age. Figure 8 below compares the counts of responses to Age Left School for the middle cohort.

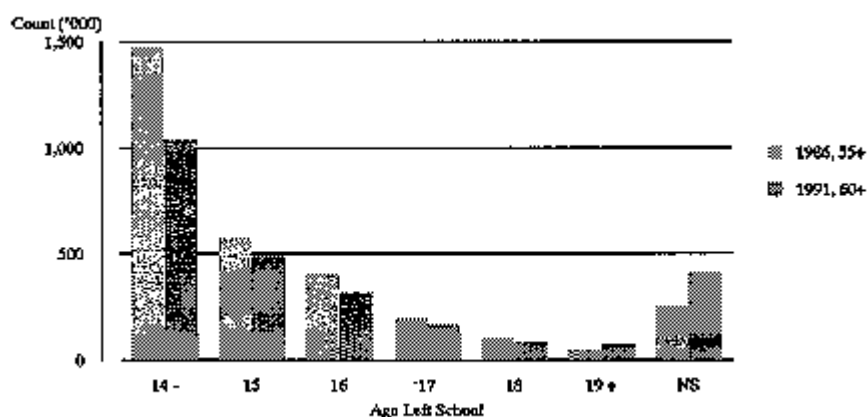
FIGURE 8: Age Left School counts, persons 40 to 54 in the 1986 Census and persons 45 to 59 in the 1991 Censuses



For the middle cohort group, the counts are generally very similar between 1986 and 1991. There are some slight changes repeating those observed for the other groups, that is a decrease in responses of 15 years or younger and an increase in responses of '19 years or older' and in Not Stated codes.

Figure 9 below contains the counts of responses for Age Left School for the oldest cohort.

FIGURE 9: Age Left School counts, persons 55+ in the 1986 Census and persons 60+ in the 1991 Censuses



As mentioned earlier, counts in the oldest cohort could be expected to decline as a result of people dying between 1986 and 1991. This was not the case for all categories however. While there was a large decrease in the number of people responding '14 years or younger' (the count of people in this category declined by 29.3 per cent compared with an overall decline of 14.9 per cent in the cohort) and slight decreases in most other categories, there was an increase in the number of people not responding to the Age Left School question.

One of the possible causes of this may be that, with rising retention rates in secondary schools, people may feel more reluctant to admit they left school younger than 14 and so not respond. In Working Paper 93/2, it was generally observed (with some exceptions) that people were less likely to mark the extreme categories on lists. While this does not appear to have affected responses to '19 years or older' it may have affected responses of '14 years or younger', where greater stigma may be felt, with people not marking any category rather than mark the lowest.

Not only do the differences between the three cohorts show the gradual increase in the general level of schooling but also different changes between the 1986 and 1991 Censuses. The large increase in responses of '19 years or older' in the youngest cohort seems to indicate that there may have been some misunderstanding of the term 'school' or that people may have given their actual age. The decrease in responses of '14 years or younger' and increase in Not Stated codes for the oldest cohort may imply that many people, particularly those with a low level of schooling, for some reason did not respond to the question on Age Left School in 1991.

Summary

Age Left School is asked in the Census as a proxy for highest level of schooling. In part, the changes between 1981 and 1991 reflect an increase in retention rates over that period. There are, however, several causes for concern regarding the quality and adequacy of the data obtained in the 1991 Census.

Between 1986 and 1991 there was a large increase in the non-response rate for the Age Left School question. This may be the result of form design changes, particularly the introduction of Optical Mark Reading, and is of concern as it indicates that many people had difficulty answering the question. Most affected were older age groups where a corresponding decline in the number of people responding '14 years or younger' was observed.

Another problem apparent in the 1991 Census data was the large number of people who gave '19 years or older' as their Age Left School. There could be several causes of this. Possibly there was some confusion over the term 'school' which is used differently between States and in other countries, sometimes including tertiary institutions. Another factor could be that, as a result of confusion over the previous question which asked age for a second time, people may have marked '19 years or older' as an indication of their actual age. Such a pattern was observed for people aged 15 to 19. Confusion resulting from both the Age Left School question and the one preceding it may also have contributed to the high non-response rate.

The adequacy of using Age Left School as a proxy for highest level of schooling was examined using National Schools Statistical Collection data. This showed that each year of schooling covered several ages and, of greater concern, that the distribution of ages varied greatly between States and Territories. While Age Left School data may be the most appropriate to obtain in a Census, its limitations should be taken into account in any analysis.

EDUCATIONAL QUALIFICATIONS

Background

Information on Educational Qualifications has been collected in every Census since 1966. In 1991, data was available on Field of Study, Level of Attainment and Year of Qualification for people with qualifications. This information is used in assessing the skill level of the labour force and enables analysis of small areas and population groups not possible with data from Surveys. It also provides more detailed information on Field of Study not available from other sources.

In the past there have been a number of problems with the data, including very high non-response rates. In 1991, substantial changes were made to the form and a new classification and coding system, the ABS Classification of Qualifications (ABSCQ) was

introduced. This section will examine the effectiveness of these changes in producing high quality data on Post-School Educational Qualifications.

Definition of a Qualification

A qualification, according to the ABSCQ, is defined as an award for attainment as a result of formal learning from an accredited post-school institution. This definition is not included on the Census form, although the question does refer to qualifications obtained 'since leaving school', and respondents are required to interpret the term 'qualification' themselves. Some assistance was available from the Hotline. The following guide-lines were included in the 1991 Census Hotline Guide:

- People who completed a course which led to a recognised certificate or higher level qualification should provide the relevant details. For example, a trade certificate gained on completion of an apprenticeship should be included.
- Do not include 'in-house' training courses such as trade union, public service or company courses.

The interpretation of 'qualification' may have had an effect on the quality of data if people include qualifications which were out of scope (although these would have been dealt with in processing) or failed to include relevant qualifications, such as trade certificates. The 1991 form was altered to encourage the inclusion of trade qualifications and the effectiveness of this will be discussed below.

Question Design

In 1986, information on Educational Qualifications was obtained from one question with several parts. This is shown below:

As will be discussed further in the following sections, non-response to this question was very high. In testing for the 1991 Census, it was noted that:

- Every attempt should be made to keep each qualification question separate. Each time questions were combined in the 1991 testing program, the codability and reliability of the questions concerned diminished. **(Census Working Paper 91/4, Report on 1991 Census Form Design Testing)**

In the 1991 Census, the question was split into five separate questions, as shown below, and additional response categories were included for the first question. The question of Year of Qualification was on a different page to the other questions.

The effectiveness of the new question design will be examined in the following sections. In particular, the change in the response rates and the impact of the additional category in Q24 'Still studying for first qualification' will be analysed.

Processing and output

Coding of Field of Study and Level of Attainment in 1991 was performed using Computer Assisted Coding (CAC). Coders would enter key words given for field of study and then select the appropriate field and then the appropriate level of qualification from those listed. The relevant ABSCQ codes were assigned by the system. This should have reduced the scope for inconsistencies between coders from 1986 when coders used a manual to find the

qualification listed and then allocated the appropriate code which covered both field of study and level of attainment.

An edit was included in 1991 to ensure that Year of Qualification was consistent with age, that is, a person could not have completed a qualification aged less than 15 years. Year of Qualification was set to Not Stated where it was inconsistent with age. No such edit was included in 1986.

In both 1986 and 1991, an Interim Final Unit Record File, the IFURF (see Section 1), was produced after codes had been allocated and basic edits performed. Among the variables deleted between the IFURF and Final Unit Record File (FURF) was the question asking whether people had a post-school qualification. The 1986 IFURF was not retained, however, data from the IFURF in 1991 is available and has been used in this paper to study the effectiveness of the question (see Appendix 6 for the incidence of combinations of responses). Three variables on Qualifications were included on the FURF. These were: Field of Study, Level of Attainment and Year of Qualification.

In the 1991 Census, qualifications that were not within the scope of the ABSCQ were given an 'Out of Scope' code and, as a part of the reformatting of the IFURF to produce the FURF, were recoded to 'No Qualification'. In 1986, these people would have been coded as 'Not recognised or inadequately described' to Field of Study/Level of Attainment. A more detailed comparison is included in Section 3.3.

Comparability with other data

Comparison with other data sources is usually a useful tool in the analysis of Census data but this was not possible for the 1991 Census Educational Qualifications data. The 1991 Census was the first time the ABSCQ was used (although it has subsequently been introduced into other ABS collections) and the use of a different coding system meant that little comparable data was available.

Some comparison with 1986 Census data was possible, although limited to broad levels (see Section 3.3.1). It was not practical to compare counts by level as analysis revealed that the change in classification to the ABSCQ, and changes in coding practice, affected all levels, even those thought to be fairly well delineated in the old classification, such as Bachelor Degree, Postgraduate Diploma and Higher Degree. Three examples affecting a large number of cases are given. Nurses who had additional qualifications in specialist areas, such as midwifery, were coded as having a Certificate in 1986, but as having a Bachelor Degree in 1991. Teaching qualifications were coded in 1986 to Diploma or Graduate Diploma on the basis of the institution attended, while in 1991 they were coded to Undergraduate or Postgraduate Diploma based on the name given for the qualification. Accountants who gave their qualification as membership of an association, for example 'Associate of the Australian Society of Accountants' or 'CPA' had their qualifications coded to Diploma in 1986 but to Bachelor Degree in 1991. The decisions to make these changes followed extensive research and consultations with professional bodies.

Data on Educational Qualifications was collected in the ABS Educational Attainment Survey, conducted in February 1991, which asked persons aged 15 to 69 years about their schooling and post-school qualifications. Comparison with this survey was attempted but there were large differences which were most likely due to the different collection and coding methodologies used.

It is hoped that comparisons with other collections will be possible for the 1996 Census as the ABSCQ is now the standard for all ABS collections which collect data on post-school educational qualifications.

Sequencing question (Q24)

Question 24, the first of the five questions asked on Educational Qualifications in 1991, was designed primarily to provide sequencing instructions which directed people without a qualification to skip the remaining qualification questions. This question also collected some information which was used in coding qualifications. Although a similar question, with fewer response categories, was asked in 1986, the 1986 IFURF has not been retained (as mentioned in Section 3.1.3) and comparable data is not available. The distribution of responses to the 1991 sequencing question is shown in Table 16 below.

TABLE 16: Responses to Question 24 a on Educational Qualifications, persons aged 15 years and older, 1991 Census

Response status	Count	Percent of Not Stated	Percent of Total
No	7,004,503		53.5
Still at school	478,301		3.7
Still studying b	414,433		3.2
Yes	4,171,475		31.9
Not Stated			
Topic non-response c	870,953	85.7	6.7
Remainder	145,688	14.3	1.1
Total Not Stated	1,016,641	100.0	7.8
Total persons 15 years and older	13,085,353		100.0
a	'Has the person obtained a trade certificate or any other educational qualification since leaving school?'		
b	Full response category was 'Still studying for first qualification'		
c	Persons who did not have a code allocated for any of Q24, Field of Study, Level of Attainment or Year of Qualification		

Source: IFURF (see 1.3)

Over 60 per cent of people stated that they did not have a qualification while almost 32 per cent stated that they did have a qualification. Over 85 per cent of people not responding to this question did not respond to any of the qualification questions. It seems likely that many of these people did not have qualifications and so saw these questions as irrelevant and skipped them. The remaining non-respondents to this question answered one or more of the succeeding qualifications questions and some of them may actually have possessed a recognised qualification.

Responses of 'Still at primary or secondary school' and 'Still studying for first qualification'

People who were studying had the choice of marking 'No' for Question 24 or marking the relevant option to indicate that they were still studying. These extra categories were included to discourage students giving details of incomplete qualifications.

There were 597,971 primary and secondary students aged 15 years and older and 1,082,560 students at other institutions. Thus, when compared with the data in Table 16 above, 80 per cent of primary and secondary students aged 15 years and older had marked the 'Still at primary or secondary school' box. Of students at Tertiary or Other institutions (not all of whom would be studying for their first qualification), 38 per cent marked the 'Still studying for first qualification' box. The effectiveness of this category will be examined further in Section 3.5.

Responses of 'Yes, trade certificate or ticket'

People who had a qualification should have chosen the appropriate response from 'Yes, trade certificate or ticket' and 'Yes, other qualification', whereas in 1986, only one response of 'Yes' was provided. The 'Yes, trade certificate or ticket' category was included to encourage people with these qualifications to give details of their qualifications following results from testing that indicated that some people with trade qualifications had responded that they had no qualification. Due to the changes in the classification of Level of Attainment, it is not possible to determine whether this change resulted in an increase of trade qualified respondents from 1986 by comparing 1986 and 1991 Census data. The Labour Classifications and Standards section has analysed responses of 'Yes, trade certificate or ticket' and feel that the introduction of the category was effective in encouraging persons with Trade Qualifications to complete the question. The information provided was also useful in coding Level of Attainment as Level could be coded, to 'Skilled Vocational Certificate' for people who only gave field information but marked this box.

Analysis by Labour Classifications and Standards section indicated that there may have been some misinterpretation of this response category. The actual responses given to write-in questions were recorded for about two per cent of the population. The resulting file is usually referred to as the '2% verbatim file' (see Section 1) and was analysed by the Labour Classifications and Standards section. Table 17 below shows the Level of Attainment of persons who marked the 'Yes, trade certificate or ticket' on this file, a total of 23,005 people.

TABLE 17: Level of Attainment, persons who marked the 'Yes, trade certificate or ticket' category, 2% Verbatim file

Level of Attainment	Percent
Skilled Vocational Certificate	76.0
Basic Vocational Certificate	8.7
Undergraduate Diploma	4.0
Bachelor Degree, Postgraduate Degree or Higher Degree	3.5
Associate Diploma	1.7

Level Unstated	3.4
Uncodable	2.7
Total	100.0

Source: 2% verbatim file

Only three quarters of those marking the 'Yes, trade certificate or ticket' category gave a qualification which was coded as 'Skilled Vocational Certificate', the category most trade qualifications are coded to (some may also have been coded to 'Associate Diploma'. Thus, it appears that the category 'Yes, trade certificate or ticket' was misunderstood by a reasonable proportion of the population and some people may have interpreted it as including certificates generally rather than trade certificates. Changes in the wording are being investigated for the 1996 Census.

Field of Study and Level of Attainment

Comparison with 1986 counts

The effectiveness of dividing the single question used in 1986 into five separate questions in 1991 can be judged by comparing the data for Field of Study and Level of Attainment from the two Censuses. However, as mentioned earlier, due to the changes in the classification and in coding practices between 1986 and 1991, only very limited comparison is possible. Table 18 below shows an overall comparison of the data. Appendix 6 explains how the table was constructed.

Note that in 1986, people who responded 'Yes' to the first part of the question but did not answer the following questions would have been classed as 'Not recognised or inadequately described'. Only people who did not give any responses were classed as 'Not Stated'. Terminology differed between 1986 and 1991. Roughly, 'Not recognised' and 'Out of scope' are equivalent, as are 'Not classifiable' and 'Inadequately described'.

TABLE 18: Responses to Educational Qualification variables, persons aged 15 years and older, excluding persons without a qualification a, 1986 and 1991 Censuses

	Counts			Percentage Distribution		
	1986	1991	Difference	1986	1991	Difference
Either Field or Level stated and adequately described	3,155,174	4,016,713	861,539	66.2	77.4	11.2
Out of scope or neither adequately described	454,324	268,016	-186,308	9.5	5.2	-4.4
Non-response b	1,154,410	903,463	-250,947	24.2	17.4	-6.8
Total	4,763,908	5,188,192	424,284	100.0	100.0	0.0

a In the 1986 Census, persons who responded 'No' or 'Still at primary or secondary school' and did not give details of a valid qualification. In the 1991 Census, persons who responded 'No', 'Still at primary or secondary school' or 'Still studying for first qualification'.

b In 1986, no responses to any of the sequencing question, Qualification name, Name of

institution or Field of study. In 1991, no response to the sequencing question (Q24) and no information on level of attainment or field of study given.

There are several limitations in the comparison.

- Persons who said they had no qualifications in the sequencing question but gave details of valid qualifications would have been coded as having qualifications in 1986 but as not having qualifications in 1991. It is assumed that the number of people affected would be very small but would have inflated the first category for 1986.
- Another assumption made for the comparison was that the code 'Not classifiable by Level, Field not specified' in 1986 was equivalent to cases where both Field and Level were coded as Inadequately Described, or Level was Not Stated and Field was Inadequately Described in 1991.
- The introduction of ABSCQ would also have had an impact on the codes allocated.

The effect of these differences on the counts should not be significant. Given these limitations on the comparison, it appears that, although the non-response rate is still high in 1991, the response rates have improved between 1986 and 1991 and so the new question design has been effective.

There has also been a reduction in the instances where neither Field nor Level were adequately described. This may be due both to the new question design encouraging people to give more complete information and to the new coding system which was more comprehensive and accurate than that used in 1986.

Non-response rates

Non-response rates should reflect the proportion of people who should have answered a particular question but didn't. However, where people did not answer any questions on a topic (referred to as 'topic non-response'), it is impossible to determine whether or not they should have answered the question. In analysis of the 1991 Census, the method of calculating response rates was changed to treat these people in a different way than in previous Censuses.

The population used in previous years and in published tables of Census counts (for example, 'Census Characteristics of Australia, 1991 Census', ABS Cat. No. 2710.0) excluded only people who were coded as not having a qualification, that is, for the 1991 Census, people who responded 'No', 'Still at primary or secondary school' or 'Still studying for first qualification' to Question 24. This population is referred to as the 'Original population' in the table below. Thus, people who did not respond to any qualifications questions, including Year of Qualification, and for whom it could not be determined whether they had a qualification, were included in the Original population.

It is likely many of these people not responding to any questions in the topic would not have responded to the qualifications questions because they appeared irrelevant (since they did not in fact have qualifications). To obtain a more realistic estimate of non-response rates for the output qualification variables, non-respondents to the topic were excluded from the population used to calculate non-response rates in 1991 (referred to as the Adjusted population in the table below).

TABLE 19: Non-response rates to Field of Study and Level of Attainment, Original a and Adjusted b populations, 1991 Census

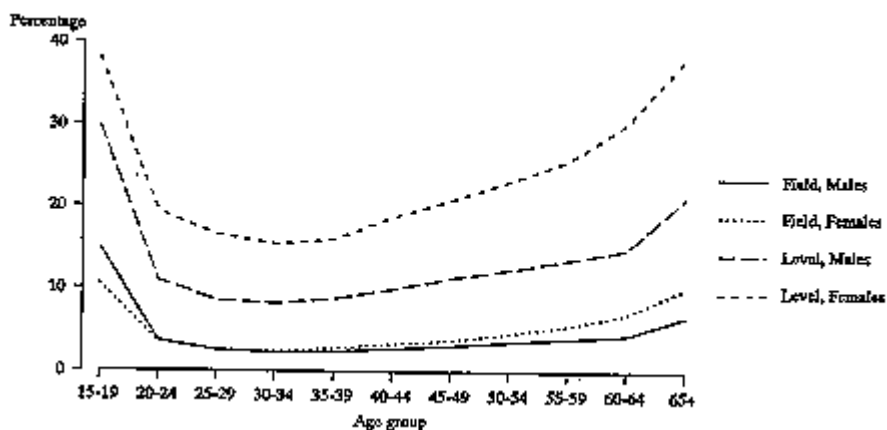
	Original a population Count	Percent	Adjusted b population Count	Percent
Field of Study				
Stated	3,973,500	78.1	3,973,500	94.2
Inadequately described	86,836	1.7	86,836	2.1
Not Stated	1,028,476	20.2	157,523	3.7
Total population	5,088,812	100.0	4,217,859	100.0
Level of Attainment				
Stated	3,456,068	67.9	3,456,068	81.9
Inadequately described	118,266	2.3	118,266	2.8
Not Stated	1,514,478	29.8	643,525	15.3
Total population	5,088,812	100.0	4,217,859	100.0
a	Excluding persons without a qualification			
b	Excluding persons without a qualification and topic non-response.			

The effect of excluding topic non-response is to reduce the non-response rates, hopefully to a level that better reflects the proportion of people who should have answered the questions but didn't. The non-response rate is much higher for Level of Attainment than for Field of Study. This is partly a result of the coding methodology. If a person gave details of Level of Attainment only, they would have been coded as having Inadequately Described (rather than Not Stated) Field of Study. However, even including the count of Inadequately Described, the proportion of people giving a codable response to Level of Attainment is lower than that for Field of Study.

The analysis of the 2% verbatim file done by the Labour Classifications and Standards section indicated that some people just described their qualification or gave Field of Study information without giving information on the Level of Attainment.

The non-response rates for Field of Study and Level of Attainment were investigated further, initially by looking at the non-response rates by age and sex, as shown in Figure 10.

FIGURE 10: Non-response rates to Field of Study and Level of Attainment, 1991 Census, by age and sex.



This graph shows that both Field of Study and Level of Attainment have similar patterns by age, with the youngest group, 15 to 19, having the highest non-response rate. The rates for the middle ages, between 20 and 44, were lower but then the rates for the older age groups gradually increased. The reasons could be connected to the apparent relevance of the question to different age groups. Persons in the 15 to 19 year age group are unlikely to have completed a qualification and therefore may see the question as irrelevant. Older people may see the Qualifications questions as irrelevant if it is many years since they completed a qualification and they no longer use it in employment or if they did not complete a qualification. The proportion of older people indicating that they had a qualification (by marking 'Yes...' for Question 24 or responding to any of the other questions) was much lower than other ages: only 20 per cent of people aged over 65 had a qualification in the 1991 Census compared to one third of persons aged 45 to 64 years and over 40 per cent of persons aged 20 to 44 years.

The other important feature of this graph is that, while the non-response rates for Field of Study are similar for both males and females, the Level of Attainment non-response rate is much higher for females. As many people responded to Field of Study but not Level of Attainment, the non-response rate for Level of Attainment was investigated further by calculating non-response rates for each Field of Study. The results are summarised in Table 20 below.

TABLE 20: Non-response rates to Level of Attainment by Field of Study, persons who responded to Field of Study, 1991 Census, by Sex

	Males		Females		Persons	
	Count of Not Stated ^a	Percentage of Total ^b	Count of Not Stated ^a	Percentage of Total ^b	Count of Not Stated ^a	Percentage of Total ^b
Business and administration	40,124	16.3	159,106	38.9	199,230	30.4
Management support services	3,034	48.0	119,068	45.5	122,102	45.6
Health	9,560	9.7	24,730	6.2	34,290	6.9
Education	6,292	5.7	26,074	8.9	32,366	8.0
Society and culture	20,411	10.6	28,042	11.7	48,453	11.2

Natural and physical sciences	12,170	9.3	11,440	15.6	23,610	11.6
Engineering	63,326	6.2	14,084	25.0	77,410	7.2
Architecture and building	17,625	5.3	1,615	15.3	19,240	5.6
Agriculture and related fields	9,103	12.8	1,767	13.9	10,870	13.0
Miscellaneous fields	18,778	12.0	15,252	13.1	34,030	12.5
Inadequately described	4,696	8.4	1,820	5.8	6,516	7.5
All stated fields of study	202,085	8.4	283,930	17.3	486,015	

a Persons who did not give information on Level of Attainment but were coded to a Field of Study

b Percentage of all people who were coded to the relevant Field of Study

The rates were by far highest, for both males and females, in the Business and Administration broad field. Within this broad field extremely high non-response rates were found for the narrow field 'Management support services'. Almost half of the people whose qualification was coded to the field 'Management support services' did not provide Level of Attainment information. This category is dominated by women and thus the high non-response to Level of Attainment for this group was a significant factor in the high non-response rate overall for women in Figure 10 above.

Analysis by the Labour Classifications and Standards section revealed that the problem was concentrated in the detailed field 'Keyboarding and Shorthand' within the narrow field 'Management support services'. This category includes as valid qualifications such as 'Associate Diploma in Secretarial Studies' and 'Certificate in Keyboarding'. Reasons for the high non-response rate for Level of Attainment may include respondents who gave details of units completed, indicating that they may not have completed a qualification, and the possibility that people gave their occupation instead of any formal qualifications they may have had.

Out of scope qualifications

Another factor in the effectiveness of the sequencing question and following questions is the type of qualifications given by respondents. The ABSCQ definition of a qualification was given in the introduction, Section 3.1.1. Some people gave qualifications, such as hobby courses, which did not fall under the definition of 'qualification'. These people were allocated a code of 'Out of Scope' in the IFURF and recoded to Not Applicable in the FURF. Of the Adjusted Population on the IFURF (excluding topic non-response), 2.3 per cent gave an out of scope qualification. This seems to indicate that, even without clarification, generally the concept of 'qualification' was well understood.

Non-response rates for Year of Qualification

Comparison with 1986 counts

Non-response to Year of Qualification was very high in 1986, however, question design has changed substantially between 1986 and 1991. In 1986, people were asked the exact year

that they completed their qualification on a line at the end of the single question on Qualification. As was shown in Section 3.1.2, in 1991 the question was split into five separate questions, with the last question on Year of Qualification on the next page and containing 7 categories for self-coding.

In 1986, tables for Year of Qualification excluded persons who were coded as Not Stated for Field of Study and Level of Attainment. If these people had been included the non-response rate would have been much higher. In Table 21 below, the non-response rate for Year of Qualification in 1991 is calculated on the same basis used in 1986, with the exception that persons who said that they did not have a qualification but gave details of a valid qualification were included in the population in 1986 but excluded in 1991. The number of persons involved should be very small.

TABLE 21: Response rates to Year of Qualification, persons who responded 'Yes...' to the sequencing question or who gave details of a field of study or level of attainment, 1986 and 1991 Censuses

		1986		1991
	Count	Percent	Count	Percent
Stated	2,804,188	77.7	4,054,718	94.6
Not Stated	805,310	22.3	229,935	5.4
Total *	3,609,498	100.0	4,284,653	100.0

* Persons who responded to the sequencing question, field of study or level of attainment. The table shows that non-response in 1991 was considerably lower than in 1986, indicating that the new question design was effective in reducing non-response.

1991 Non-response rates

Table 22 is equivalent to Table 19 in the Section 3.3.2 and shows the non-response rates for 1991 calculated using both the original population and the adjusted population.

TABLE 22: Response rates to Year of Qualification, Original a and Adjusted b populations, 1991 Census

		Original a population Percent		Adjusted b population Percent
	Count		Count	
Stated	3,998,325	78.6	3,998,327	94.8
Not Stated	1,090,485	21.4	219,532	5.2
Total	5,088,810	100.0	4,217,857	100.0

a Excluding persons without a qualification

b Excluding persons without a qualification or not responding to any of the qualification questions.

The non-response rate for Year of Qualification is not much higher than that for Field of Study and does not have the apparent problems seen for Level of Attainment. Thus, the question design seems to be effective and the concept of Year of Qualification easily

understood.

Commission error

One problem resulting from the new form design was over-response, or commission error. In 1991, the question on Year of Qualification was at the top of a page, after the other Qualification questions. Experience has shown that people are more likely to respond to questions at the top of a page than at other positions and it appears that many people did respond to this question even though they had indicated that they did not have a qualification. Responses given where none were required are referred to as 'commission error'. Table 23 below shows the extent of commission error for Year of Qualification according to category of response to the sequencing question, Question 24.

TABLE 23: Responses to Year of Qualification by Response to Question 24 a, persons without a qualification, 1991 Census

Year of Qualification	Response to sequencing question (Q24)							
			Still at school		Still studying		Total stating no qualification	
	No.	%	No.	%	No.	%	No.	%
Year not stated	6,785,991	96.9	474,828	99.3	399,637	96.4	7,660,456	97.0
Year stated (Commission error)	218,512	3.1	3,473	0.7	14,796	3.6	236,781	3.0
Persons without a qualification	7,004,503	100.0	478,301	100.0	414,433	100.0	7,897,237	100.0

a 'Has the person obtained a trade certificate or any other educational qualification since leaving school?'

Source: IFURF (see Section 3.1.3)

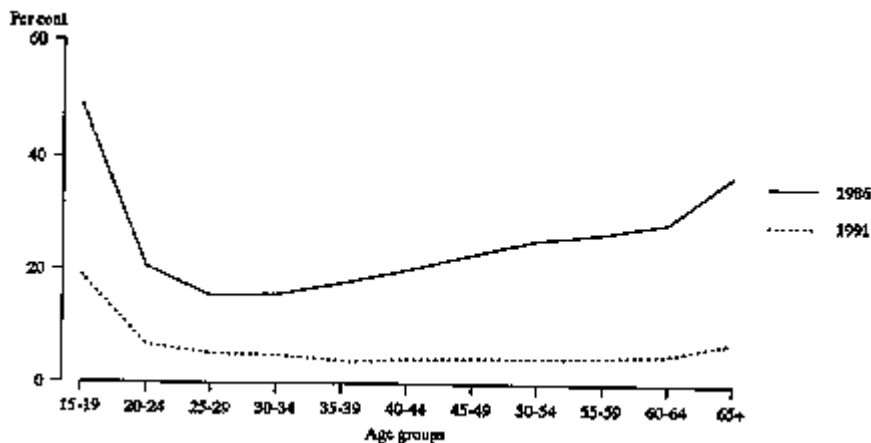
These errors would have been corrected in the FURF where the response to Year of Qualification would have been set to Not Applicable. The level is reasonably low, as would be expected since the question is unlikely to attract a response from many people without qualifications.

Effectiveness of the 'Still studying for first qualification' response category

In 1991 an additional response category of 'Still studying for first qualification.' was introduced for Question 24 'Has the person obtained a trade certificate or any other educational qualification since leaving school?'. This was done as a result of concerns that many people aged 15 to 24 who were studying for their first qualification may have answered the series of questions on qualifications in the 1986 Census incorrectly, giving details of incomplete qualifications and usually not responding to Year of Qualification.

The graph on the following page indicates that the inclusion of the additional response category may have lessened the problems observed for 1986. It shows that the non-response rate for Year of Qualification was lower in 1991 than 1986.

FIGURE 11: Non-response rates for Year of Qualification by age, 1986 and 1991 Censuses



The non-response rate for Year of Qualification reduced in all age groups between 1986 and 1991. In particular, the non-response rate for persons aged 15-19, although still high, has fallen considerably. This supports the theory that fewer persons still studying for their first qualification attempted to give responses to the qualifications questions. Table 24 below shows this even more clearly. The categories used are the same as those used in Table 18, which are comparable between 1986 and 1991.

TABLE 24: Distribution of responses to Qualifications questions, persons aged 15 to 24 years attending a Tertiary institution (TAFE, University or CAE), 1986 and 1991 Censuses

	Either Field or Level stated and adequately described %	Out of scope or neither adequately described %	Non-response a %	No Qualification b %	Total %
Persons aged 15 - 19 years					
1986	5.9	11.8	15.8	66.5	100.00
1991	3.3	1.2	2.4	93.0	100.00
Persons aged 20 - 24 years					
1986	26.4	7.5	11.4	54.7	100.00
1991	22.0	2.0	2.1	74.0	100.00

a In the 1986 Census, persons who responded 'No' or 'Still at primary or secondary school' and did not give details of a valid qualification. In the 1991 Census, persons who responded 'No', 'Still at primary or secondary school' or 'Still studying for first qualification'.

b In 1986, no responses to any of the sequencing question, Qualification name, Name of

institution or Field of study. In 1991, no response to the sequencing question (Q24) and no information on level of attainment or field of study given.

Although the number of people attending a TAFE, University/CAE who gave details of a qualification fell only slightly between 1986 and 1991, the proportions giving invalid or inadequately described qualification or not responding fell considerably, while the number responding that they did not have a qualification. This indicates that more people who have not yet completed their first qualification responded correctly rather than trying to give details of incomplete qualifications or skipping the question completely. Again, the new question design used in the 1991 Census appears to have been effective in improving the quality of responses to the Qualifications questions by eliminating many responses relating to courses not yet completed.

Summary

The changes introduced in the 1991 Census, both the new question design and the new classification and coding system, appear to have improved the quality of Qualifications data.

The new separation of the single multi-part question used in the 1986 Census into five questions appears to have reduced non-response rates, particularly for Field of Study and Year of Qualification. However, the non-response rate for Level of Attainment is still high. This is concentrated in some fields, in particular in Management support services.

The introduction of new categories in the initial sequencing question also seems to have had a positive effect by discouraging persons who have not yet completed their qualifications from answering the questions and by encouraging persons with trade qualifications to answer the questions.

About this Release

ABOUT THIS RELEASE

The aim of this paper is to provide users with a detailed assessment of the benefits and limitations of Census data for each of the three education topics contained in the 1991 Census. Each topic is analysed separately looking, in particular, at factors such as variation in counts between censuses, changes in form design, non-response and comparisons with other education data sources to evaluate data quality.
